6-1-2012

The tegula tango: A coevolutionary dance of interacting, positively selected sperm and egg proteins

Michael E. Hellberg
Louisiana State University

Alice B. Dennis
Louisiana State University

Patricia Arbour-Reily
Louisiana State University

Jan E. Aagaard
University of Washington, Seattle

Willie J. Swanson
University of Washington, Seattle

Recommended Citation

Hellberg, M., Dennis, A., Arbour-Reily, P., Aagaard, J., & Swanson, W. (2012). The tegula tango: A coevolutionary dance of interacting, positively selected sperm and egg proteins. *Evolution, 66* (6), 1681-1694. https://doi.org/10.1111/j.1558-5646.2011.01530.x

doi:10.1111/j.1558-5646.2011.01530.x

THE *TEGULA* TANGO: A COEVOLUTIONARY DANCE OF INTERACTING, POSITIVELY SELECTED SPERM AND EGG PROTEINS

Michael E. Hellberg,[1,2] Alice B. Dennis,[1,3] Patricia Arbour-Reily,[1] Jan E. Aagaard,[4] and Willie J. Swanson[4]

[1]*Department of Biological Sciences, Louisiana State University, Baton Rouge, Louisiana 70803*

[2]*E-mail: mhellbe@lsu.edu*

[3]*Landcare Research, Private Bag 92170, Auckland 1142, New Zealand*

[4]*Department of Genome Sciences, University of Washington, Seattle, Washington 98195–5065*

Received August 8, 2011
Accepted November 11, 2011

Reproductive proteins commonly show signs of rapid divergence driven by positive selection. The mechanisms driving these changes have remained ambiguous in part because interacting male and female proteins have rarely been examined. We isolate an egg protein the vitelline envelope receptor for lysin (VERL) from *Tegula*, a genus of free-spawning marine snails. Like VERL from abalone, *Tegula* VERL is a major component of the VE surrounding the egg, includes a conserved zona pellucida (ZP) domain at its C-terminus, and possesses a unique, negatively charged domain of about 150 amino acids implicated in interactions with the positively charged lysin. Unlike for abalone VERL, where this unique VERL domain occurs in a tandem array of 22 repeats, *Tegula* VERL has just one such domain. Interspecific comparisons show that both lysin and the VERL domain diverge via positive selection, whereas the ZP domain evolves neutrally. Rates of nonsynonymous substitution are correlated between lysin and the VERL domain, consistent with sexual antagonism, although lineage-specific effects, perhaps owing to different ecologies, may alter the relative evolutionary rates of sperm- and egg-borne proteins.

KEY WORDS: **Gamete recognition, gastropod, lysin, speciation, VERL, zona pellucida.**

Proteins involved with sexual reproduction commonly show signs of rapid divergence promoted by positive selection (Swanson and Vacquier 2002). Evidence of rapid divergence was first noted for proteins on sperm and in seminal fluid (accessory proteins in *Drosophila*: Tsaur and Wu 1997; Aguadé 1999; lysin in gastropods: Lee et al. 1995; Hellberg and Vacquier 1999; lysin M7 in bivalves: Riginos and McDonald 2003; Springer and Crespi 2007; bindin in sea urchins: Metz and Palumbi 1996; seminal proteins in mammals: Clark and Swanson 2005), but subsequent work has found rapid change in female reproductive proteins as well (*Drosophila*: Swanson et al. 2004; Kelleher et al. 2007; gastropods: Galindo et al. 2003; mammals: Swanson et al. 2001). Rates of change in sex proteins can be startlingly fast: the major acrosomal protein (TMAP) from sperm from the marine snails *Tegula regina* and *T. montereyi* is nearly 60% divergent at the amino acid level (Hellberg et al. 2000) even though mitochondrial DNA (COI) sequences are less than 6% divergent for this species pair (Hellberg 1998). Understanding the evolutionary forces driving such rapid divergence would not only reveal the forces behind selection at its most extreme but, because the functions of some of these proteins should be sufficient to confer reproduction isolation (and indeed sequence divergence is correlated with reproductive compatibility: McCartney and Lessios 2004; Zigler et al. 2005), might reveal mechanisms contributing to the formation of new species.

That natural selection promotes the rapid divergence of sex proteins is evidenced by high rates of nonsynonymous (amino acid changing) nucleotide substitution (d*N*) relative to rates of

synonymous (silent) change (d*S*). But what forces underlie this positive selection on sex proteins? Several hypotheses have been put forward (Howard 1999). Forces external to the sperm and egg may act on one or both interacting protein partners. Egg coats, for example, may function not only in recognizing conspecific sperm, but also in defending the fertilized egg against microbes (Vacquier et al. 1997; Turner and Hoekstra 2008) or adapting the egg to substrate conditions (Jagadeeshan and Singh 2007). An egg protein's evolutionary response to pathogenic pressure could force an interacting sperm protein to follow along. Sperm-recognition proteins might also diverge to avoid fertilizing co-occurring heterospecific sperm when such unions produce offspring with low fitness (Metz et al. 1998a; Geyer and Palumbi 2003; Springer and Crespi 2007 but see Geyer and Lessios 2009). Internal conflict within a single genome can drive a reciprocal escalation between the male and female sexual characters if males and females assume different strategies to maximize their reproductive success (Rice 1996; Rice and Holland 1997; Clark et al. 1999). Males, for example, may try to maximize their fertilizations, even at the cost of reducing female longevity (in internal fertilizers) or wasting female eggs due to polyspermy (in external fertilizers).

Under such sexual antagonism, the associations between male and female sex proteins are expected to be close. Palumbi (1999) found the success of sperm with a particular genotype for bindin (a sea urchin protein that binds to the egg's surface) was positively associated with the *female's* bindin genotype, even though bindin is not expressed on the egg. Subsequent work by Levitan and his colleagues (Levitan and Ferrell 2006; Levitan et al. 2007; Levitan and Stapper 2010) found similar associations between bindin genotypes of sperm and egg. Furthermore, they found particular genotypes did best under usual conditions of sperm limitation, whereas others did best when gamete concentrations led to polyspermy and sexual antagonism. These results substantiate the impact variation at a single sex protein coding locus can have and the potential for coevolutionary escalation between male and female sex proteins. The most direct way to assess whether sexual antagonism drives rapid change in sex proteins would be to compare interacting proteins from sperm and egg. Although interacting male and female protein pairs have been identified in sea urchins (Kamei and Glabe 2003), to date coevolutionary comparison has been made for just one pair: lysin and vitelline envelope receptor for lysin (VERL).

Lysin is expressed in the acrosome of sperm from vetigastropods, including *Haliotis* (abalone) and *Tegula*, snail genera that diverged about 250 Mya (Tracey et al. 1993). Upon contact with the VE, a glycoproteinaceous layer that surrounds the egg, lysin is released and proceeds to dissolve a hole in the VE via a nonenzymatic mechanism (Haino-Fukushima 1974; Lewis et al. 1982). This VE dissolution is highly species-specific in both *Tegula* (Haino 1971; Haino-Fukushima et al. 1999; Hellberg and Vacquier 1999) and *Haliotis* (Vacquier et al. 1990). Positive selection promotes the rapid interspecific divergence of lysin in both abalone (Lee and Vacquier 1995) and *Tegula* (Hellberg and Vacquier 1999), although variation within species is very low (Metz et al. 1998b; Clark et al. 2009). The shared localization, function, and structure of the lysins of abalone and *Tegula* suggest they are orthologs (Hellberg and Vacquier 1999).

The egg-borne protein that interacts with lysin, VERL, has been characterized and cloned in abalone (Swanson and Vacquier 1997; 1998; Galindo et al. 2002). The bulk of this large protein (>2 million Daltons [MDa], about half of that carbohydrate, and a cDNA of 11,116 bp) is composed of 22 repeats of a unique domain about 150 amino acids in length (Galindo et al. 2002). Most of the 22 VERL repeats evolve via concerted evolution, such that repeats within a species are more similar to each other than to any in different species (Swanson and Vacquier 1998). However, the first two VERL repeats do not undergo concerted evolution and instead diverge via positive selection (Galindo et al. 2003). Biochemically deduced stoichiometry (Swanson and Vacquier 1997), combined with the known lysin structure (Shaw et al. 1993; Kresge et al. 2000), suggest that lysin acts via a two-step mechanism (Kresge et al. 2001). In the first step (primary recognition), dimerized lysin binds to VERL in a species-specific manner. This step separates the lysin dimer, whose monomers then, in the second step (secondary recognition dissolution), bind to nearby VERL repeats, thereby breaking the hydrogen bonds that join them and dissolving the VE.

Lysin and VERL may meet the criteria laid out by Coyne and Orr (2004, pp. 244–245) for gamete recognition proteins coevolving via sexual antagonism. Free-spawning marine invertebrates often spawn at the same time (e.g., Stekoll and Shirley 1993) and have blocks to polyspermy (Jaffe and Gould 1985; Stephano 1992). Both lysin and some VERL repeats diverge via positive selection, and these rates (as measured by the ratio of rates of nonsynonymous to synonymous substitution, d*N*/d*S*) are correlated (Clark et al. 2009). The repetition of VERL domains, however, makes further inferences difficult, because although only the first two appear involved in recognition and these have broken free from concerted evolution (Galindo et al. 2003), these still provide two moving targets against which a single lysin must contend.

Here, we describe the isolation of a VERL ortholog from *Tegula* which contains just a single VERL domain. This simple structure allows direct comparison of the relative rates of change to interacting male and female sex proteins. Results show that these proteins have diverged at high rates via positive selection and that these rates are correlated for lysin and the part of VERL implicated in sperm recognition but not for the zona pellucida (ZP) domain thought to play a structural role. Relative rates of divergence for the sperm and egg proteins appear to vary, perhaps due to lineage or ecological factors.

Methods

CLONING AND SEQUENCING OF VERL cDNA

Eggs were obtained from *Tegula funebralis* by cracking shells with a hammer, removing gonads with a razor, and shaking gametes free into chilled seawater. Eggs were washed by settling three times in 1 mg/mL BSA seawater. VEs were isolated by multiple rounds of Dounce (Teflon-glass) homogenization followed by gentle centrifugation (Lewis et al. 1982).

VEs were solubilized by acid dissolution then underwent centrifugation after being returned to pH 7.8 seawater (Swanson and Vacquier 1997). VE proteins were separated on a 2.5% polyacrylamide gel (Swanson and Vacquier 1997). The largest protein on these gels was > 1 MDa and highly abundant. Because abalone VERL is also very large and constitutes about 30% of soluble VE material from *Haliotis rufescens* (Swanson and Vacquier 1997), we reasoned these large VE proteins might be orthologous and submitted the excised protein to the PAN facility at Stanford University for amino acid sequencing.

Peptide sequences from the peptide fragments of the candidate VERL protein were used to design degenerate polymerase chain reaction (PCR) primers. Forward primer TV5 (GARATHGCNACNCAYCARGGNGG, matching AA sequence IATHQGG) and reverse primer TV8 (TTNGTYTCDATDAT NGG, matching AA sequence PIIETN) amplified an 88-bp fragment that matched the intervening amino acid sequence from the fragment for which these were designed. An exact match reverse primer designed from this sequence, TV16 (TCACACCC TCCTCTCAGAATAG), in combination with a degenerate forward primer TV1 (CARGAYACNGTNAAYTTYTAY, matching peptide sequence QDTVNFYL) made to match another peptide fragment amplified a 473-bp fragment whose inferred amino acid sequence matched the fragments and the ZP domain shared by abalone VERL and two VE proteins previously isolated from another *Tegula* species (Haino-Fukushima et al. 2000).

Total RNA was isolated from the ripe ovaries of three *T. funebralis* individuals using standard guanidinium isothiocynate/CsCl techniques (Chomczynski and Sacchi 1987). From 478 μg total RNA, 12 μg mRNA was isolated using the Micro PolyA Pure Kit (Ambion, Foster City, CA). This mRNA was used to construct a cDNA library (Stratagene Zap-cDNA kit, Cedar Creek, TX).

Plaques from the ovary cDNA library were screened using a ^{32}P-labeled TV1/TV16 fragment. We recovered several positive clones, the five largest about 1.2 Kb each. The 5′ end of all of these included a repeat array (see Results) that the RT had apparently been unable to progress through during cDNA library construction. Multiple attempts to progress beyond these repeats using either reverse primers from the 3′ end of the cDNA paired with a forward primer from the plasmid vector or inverse PCR failed.

To advance beyond the repeats, we made a genomic DNA library. gDNA was isolated from *T. funebralis* using the Qiamp DNA isolation kit (Qiagen, Germantown, MD). Southern digests (again probed using radio-labeled TV1/TV16 amplicons) were performed using several restriction enzymes to insure DNA digests containing the putative VERL sequence would be large, and thus likely to extend beyond the repeats. A total of 100 μg of the *T. funebralis* gDNA was restricted using EcoRI (New England Biolabs, Beverly, MA) and run out on a 0.9% TAE gel. DNA in the 3–6 Kb region was excised from the gel and purified (Qiagen Qiaex Gel extract II kit). Recovered DNA was ligated to 1 μg Lambda Zap vector arms from the Stratagene Lambda ZAP II kit using T4 DNA ligase (NEB), with an extra 0.5 μL of 10 mM rATP added to the 50 μL reaction. After overnight ligation at 4°C, the vector was packaged using Lambda Gigapack III Gold packaging extracts (Stratagene). Plaques were screened as before. The largest of the positive clones (about 3.5 Kb) included both the previously sequenced 3′ end from the cDNA library and extended about 350 bp beyond the repeat region.

The 5′ RACE was used to complete the cDNA sequencing. The 5′ RACE cDNAs were sequenced from 1 μg of total RNA from *T. funebralis* using the SMART RACE cDNA amplification kit (Clontech, Madison, WI). In separate reaction, two reverse primers (TegFuVERL-PEAr, GATAAGCAAGACCCGCGAAAT GGTCG, and TegFuVERL-MQVr, CACTTGCATATCGCTTA-CAATTGGTGCG) were used in combination with the 5′ CDS forward primer from the kit. Resulting products were cloned using the TOPO TA Cloning Kit for Sequencing (Invitrogen). Degenerate PCR primers were designed based on the alignment of the resulting *T. funebralis* sequence with the VERL domain of *Haliotis rufescens*. One of the reverse primers (drTV5CYVr, GGTGGGCAACCACATTCNACRTAGCA) was used to obtain 5′ sequence of the putative VERL cDNA from a second *Tegula* species, *T. xanthostigma*.

OBTAINING VERL AND LYSIN cDNA SEQUENCES FROM MULTIPLE SPECIES

Gonad samples were obtained from 10 species of *Tegula* (Table 1) that mtDNA data suggest constitute three clades of three closely related species plus a more distant outgroup (Hellberg 1998). One of these 10 species (*T. lischkei*) was not included in a previous phylogenetic study of this genus, but has sometimes been synonymized with a species that was (*T. argyrostoma*). Mitochondrial COI sequence (JQ182426, obtained using the primers of Folmer et al. 1994) confirms that it is closer to the (*T. argyrostomait* + *T. xanthostigma*) sister group than to any other species analyzed here.

Total RNA was isolated from ripe ovaries (as for the cDNA library construction) or using Trizol reagent (Invitrogen). First-strand cDNA was synthesized using oligo-dT

Table 1. *Tegula* species used in this study.

Species	Collection locality	Collector
T. aureotincta	San Diego, California, USA	MEH
T. brunnea	Pacific Grove, California, USA	L. Tomanek
T. montereyi	Pacific Grove, California, USA	MEH
T. regina	San Diego, California, USA	R. McConnaughey and V. Vacquier
T. rusticus	Kagoshima, Kyushu, Japan	Y. Tomoko
T. xanthostigma	Uchiura, Honshu, Japan	Y. Hirano
T. lischkei	Uchiura, Honshu, Japan	Y. Hirano
T. funebralis	Cape Arago, Oregon, USA	P. Marko
T. gallina	Pta. Eugenia, Baja Calif., Mexico	P. Fenberg
T. rugosa	Pto. Peñasco, Sonora, Mexico	R. Grosberg

as a primer. This single-stranded cDNA was used as template for amplification of the putative VERL cDNA in three overlapping parts. The VERL domain was amplified using primers TegVERL-KWTf (GCGCCATGAARTGGAC) and drTV5CYVr with an annealing temperature of 55°C. The short repeats in the middle of the cDNA were flanked by forward primer TV35 (TTAATCGAGTGGAGCGATGG) and reverse primer closeSPTTrev (CCACAACGAGTAACCACT TCAA). The ZP domain at the end of the cDNA was amplified using primers TegVERL-RCGf (GGATTGTTGAAGTGG TTACTCGTTGTGG) and TegChlorVERL-3UTR (GAATCT TAGCTACATCACTGCG).

Lysin cDNA sequences for *T. aureotincta*, *T. brunnea*, and *T. funebralis* have been published previously (Hellberg and Vacquier 1999). RNA was isolated from testes as from ovaries. After first-strand synthesis of cDNA using oligo-dT, lysin cDNA was amplified using primers TegLys5Sig2 (TACYAGAATGAAAG GTGCNGTNSTGTG) and Tpfi3end (TGAAGCCCTAAATATA CATTTATT).

Multiple attempts to secure RNA from the testes of *T. rugosa* failed, so we tried to amplify lysin exons from genomic DNA using primers based on its close relatives *T. funebralis* and *T. gallina*. Forward primer TegLys5Sig4 (GCATCATACYAGAAT GAAAGGTGC) and reverse primer TfLI1RevA (GAACATAC CTGCTTTTATCAGCCC) amplified exon 1 (and intron 1) from *T. rugosa*.

PHYLOGENETIC ANALYSES

The VERL and ZP domains of the putative VERL contained no indels (save a 3-bp in-frame deletion in the ZP domain of *T. rusticus*) and were aligned unambiguously. Inferred lysin amino acid sequences included a few in-frame 1- or 2-residue indels near the carboxy-terminus; these were aligned by hand to try and maintain residue similarity and minimize indel events, then converted back to nucleotides for phylogenetic analysis.

VERL cDNA nucleotide sequences were subdivided into three nonexclusive datasets for analysis: the mature VERL domain, the ZP domain, and a concatenation of these two. Appropriate models for phylogenetic analysis were selected using JModel test (Posada 2008) based on the Akaike or Bayesian information criteria for the Bayesian and likelihood analyses, respectively (Huelsenbeck and Rannala 2004; Posada and Buckley 2004). We excluded the possibility of using a model with both a gamma-distributed rates and a category for invariant sites because the two are not independent. The selected model (TrN + G) could not be implemented in MrBayes (version 3.1.2, Ronquist and Huelsenbeck 2003), so the GTR + G model (which should converge with the former) was used for a Bayesian phylogenetic analysis. Analyses were run for one million generations, with sampling every 1000 generations. Samples before convergence (determined by inspection in Tracer version 1.5, Rambaut and Drummond 2010) were discarded as burnin. Maximum-likelihood (ML) trees and bootstrap support values were determined using GARLI 0.95.1 (Zwickl 2006), again using gamma-distributed rates but without an invariant sites category.

TESTS FOR POSITIVE SELECTION

Positive selection leaves its signature on gene coding regions of DNA as a ratio of nonsynonymous substitutions per nonsynonymous site to synonymous substitutions per synonymous site (dN/dS, or ω) being greater than 1. To test for positive selection acting on lysin and the two nonrepeat domains of VERL, we used the codeml program in PAML 4.2a (Yang 2007). Pairwise measures of dN/dS for all combinations of species were obtained using the measure of Nei and Gojobori (1986) as calculated using MEGA5 (Tamura et al. 2011). An overall estimate for dN/dS from assuming the same ratio at all codons and on all branches was obtained using codeml model M0. Two likelihood ratio tests were used to test for the significance of selection. First, nested models M7 and M8 were compared (Yang et al. 2000). M7 constrains evolution to be neutral; ω is allowed to take on beta-distributed values between 0 and 1. Model M8 includes an additional rate class where ω can exceed 1. Model M8 was also compared to model M8a (Swanson et al. 2003), in which an additional rate class is set to $\omega = 1$. The significance of selection was determined by comparing twice the difference likelihood values of M8 and its alternative to a chi-squared table of critical values with

degrees of freedom equal to the differences in the number of model parameters.

Selected changes at individual residues were analyzed in two ways. First, individual residues evolving under positive selection were identified using a Bayes empirical Bayes (BEB) approach (Yang et al. 2005). Second, changes along particular branches were inferred using ancestral sequence reconstruction (Yang et al. 1995) as implemented in codeml.

An ML model was used to test the degree to which positive selection was correlated for lysin and VERL (Clark et al. 2009). A free model was created by partitioning the data into lysin and VERL and assigning each a tree topology and codon model. Values of dN and dS were estimated using the Goldman and Yang (1994) codon model (GY94), which places constraints on the linear model where slope and y-intercept are global parameters:

dN/dS(VERL, branch i) = slope × dN/dS(lysin, branch i) + y-intercept.

This constraint was replicated for all corresponding branches (i). A nested null model was defined by setting the slope to zero. Comparing these nested models tests whether the slope parameter is significantly nonzero or dN/dS ratios are correlated. There is one degree of freedom between the linear and null models. A P-value was obtained using a likelihood ratio test. Both VERL and lysin were analyzed as the dependent variable of the linear relationship. We estimated our likelihood using custom scripts written for HyPhy (Hypothesis Testing Using Phylogenies) version 0.9920060106beta for Macintosh OSX (Pond et al. 2005). Tree distances were used as initial parameter values, and the optimization used settings at a precision of 1×10^{-5} and persistence of "very high."

Results

THE PUTATIVE VERL cDNA SEQUENCE FROM *T. FUNEBRALIS*

A single large (>1 MDa) protein constituted about 90% of the soluble VE material from *T. funebralis*. Three peptide sequences were obtained from this large *Tegula* VERL candidate: SQDTVNFYLPVYYDPSK, EIATHQGGTSSSAITLQVVDVL GRPIIETN, and NEQSLVSLK. The full-length candidate VERL cDNA isolated from *T. funebralis* using a combined library screening and 5'RACE approach was 2331 bp in length (GenBank accession number JQ182423). One of the three directly sequenced peptides (the shortest) occurred near the start of the inferred amino acid sequence, 22 amino acid residues after a recognizable start methionine. The SQD peptide begins near the start of the ZP domain (see below) and the EIA peptide a little further on, also in the ZP domain. The mature protein had a predicted length of 716 residues and consisted of three different parts (Fig. 1).

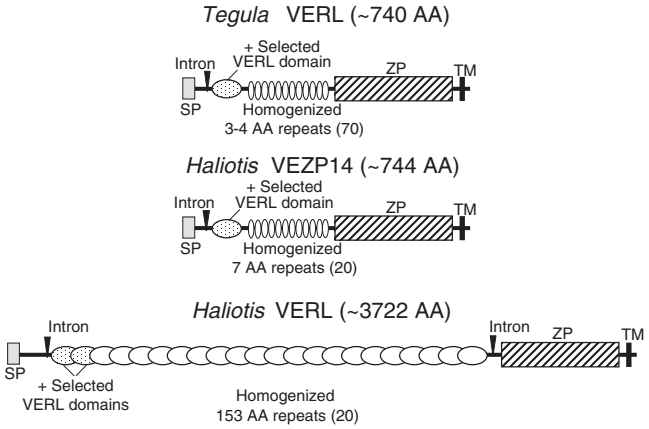

Figure 1. Gene structure supports the orthology of the putative VERL from *Tegula funebralis* (top) and two paralogous proteins from the abalone *Haliotis rufescens*, VERL and VEZP14.

The first portion of the mature protein (114 residues) aligned to a unique domain found in two VE proteins from abalone: VERL and VEZP14 (VE ZP 14), a VERL paralog that may be the receptor for a rapidly evolving paralog of lysin (Aagaard et al. 2010). As in these VERL domains, the 5′ part of the *Tegula* VERL candidate is negatively charged, with an isoelectric point of 4.28 for *T. funebralis*, in strong contrast to the positively charged lysin (pI = 11.19 for T. *funebralis*). The sole intron in the *T. funebralis* gene occurs just before the VERL domain, a position shared with the two abalone proteins (Fig. 1, see also Aagaard et al. 2010). Beginning immediately after this residue, the putative *Tegula* VERL protein aligns well with its *Haliotis* ortholog (Fig. 2). Although amino acid identity is low, all four cysteines and the single tryptophan that are conserved among the selected first and second VERL repeats and the later homogenized ones are also conserved in the putative *Tegula* VERL repeat.

The middle of the protein (244 residues) was composed of a homogenized array of three or four -residues, mainly SPT and SPTT, tandemly repeated 70 times. This region was like the middle section of the abalone protein VEZP14 in possessing short repeats (TTTTTTP for VEZP14 from *H. rufescens*, Aagaard et al. 2010) rich in threonines, whose hydroxyl groups (like those of serine) can be glycosylated or act in hydrogen bonds.

The remainder of the mature protein includes the conserved cysteine backbone characteristic of ZP proteins. These (Jovine et al. 2005) were first identified from the ZP surrounding mammalian eggs, but also occur in the VEs of *Haliotis* (Aagaard et al. 2006, 2010) and *Tegula* (Haino-Fukushima et al. 2000). Previous phylogenetic analysis of the amino acid sequence from this *T. funebralis* ZP domain has placed it in a clade with abalone VERL and VEZP14, the two proteins with VERL domains (Aagaard et al. 2010). As in *Haliotis* VERL and VEZP14, a trans-membrane

```
TfunVERL   EVLALGNRINGMNITKNCSDDFDELSEFNFYFDYTVLNVTVVNISLACLGSTFDSMRV
HrufVERL1  DADTPDPRVLSLDLTLVCSDD-KSKQATLISYPVTFKGHVIKDMQIFCKNGWMQMTRG
HrufVERL2      SHNQSKLIDWDVFCSQN-ENIPAKFISRLVAPKCLAVEKMDVDCSNGLVPITHE
HrufVERL3      SHNQSKLIDWDVYCSQD-ESIPAKFISRLVTSKDQALEKTEINCSNGLVPITQE

Tfun  EDDHFAGLAYHPEASSPVENKCYF-APIVSDMQVKHILY---FLIEWSDG------KLCYVECGC
Hr1   RGINMIRIHYPQTYTSVVPGACVFRGPYSVPTNDSIEMYNVSVALLWSDGTPTYESLECNVTKSQ
Hr2   HGFNMMLIQYTRNKLLDSPGMCVFWGPYSVPKNDTVVLYTVTARLKWSEGPPTDLSIQCYMPKSP
Hr3   FGINMMLIQYTRNELLDSPGMCVFWGPYSVPKNDTVVLYTVTARLKWSEGPPTNLSIQCYMPKSP

Tfun  P--PPSR P----------ASPTSPA
Hr1   ASNAPEPK----------ASPTSSTPQPEAA
Hr2   D--APKP ESCLSSPPEPEASPSSNAPEPETYPTSSAP
Hr3   V--APKP E----------TGPTSNAPEPETYPTSSAPEKVSSDQPAP
```

Figure 2. Amino acid alignment of amino-terminal region of putative VERL from *Tegula funebralis* with VERL from *Haliotis rufescens*. The alignment between the *T. funebralis* sequence and the first *H. rufescens* repeat begins where a shared intron occurs. Conserved residues are shaded.

domain immediately follows the ZP domain of the *Tegula* protein, with a furin cleavage site following in turn. Unlike the VERL domain or lysin, the ZP domain is weakly charged (pI = 6.24).

GENE TREES FOR LYSIN AND VERL

The topologies of the phylogenetic trees (Fig. 3) based on lysin (GenBank accession numbers JQ182410-JQ182415) and on the three partitions of VERL (JQ182416-JQ182425) agree generally on basal relationships. Each of the three three-species radiations supported by mtDNA data (Hellberg 1998) receive significant support (BPP $\geq$ 0.99, ML bootstraps $\geq$ 0.8). The sister relationships between the subtidal Californian species (*T. brunnea*, *T. montereyi*, and *T. regina*) and the Japanese species (*T. rusticus*, *T. xanthostigma*, and *T. lischkei*) found in mtDNA are well supported by the ZP domain data, less so by the two other VERL partitions, and not at all for the lysin data. Relationships within the North American intertidal clade of *T. funebralis*, *T. gallina*, and *T. rugosa* are also in agreement with the mtDNA tree and

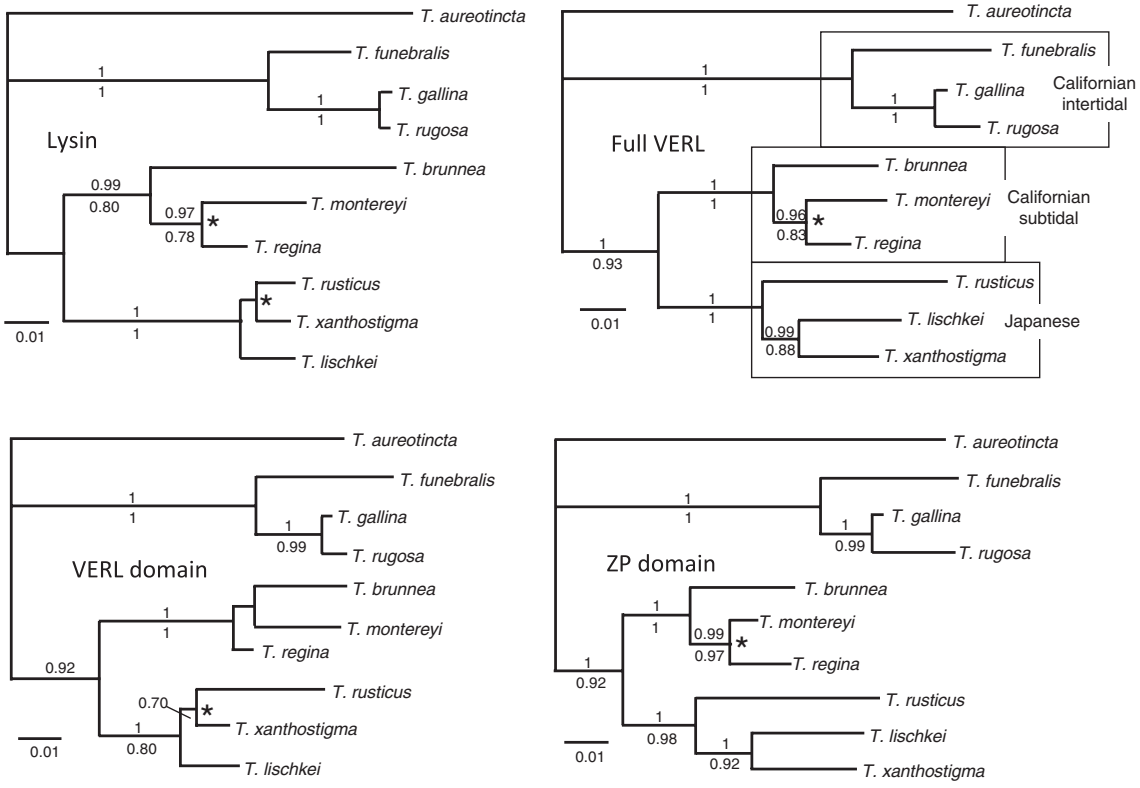

Figure 3. Gene trees for the sperm protein lysin and its egg-borne receptor VERL. Trees for VERL based on just the N-terminal region of the protein (VERL domain), the C-terminal region (ZP domain), and the two of these concatinated (full VERL). Maximum likelihood trees are shown, with bootstrap support values above each branch and Bayesian posterior probabilities above. Sister groups in conflict with the mitochondrial topology of Hellberg (1998) are marked with an asterisk. Clades referred to in text are boxed.

strongly supported (BPP = 1, ML bootstrap ≥ 0.99) for lysin and the three VERL partitions.

Relationships among the three species in the subtidal California clade and the Japanese clade varied among mtDNA (Hellberg 1998), lysin, and the VERL partitions (Fig. 3). Support for a sister relationship between *T. montereyi* and *T. regina* is high for lysin (BPP = 0.97, ML bootstrap = 0.78) and more so for the ZP domain of VERL (BPP = 0.99, ML bootstrap = 0.97), although this conflicts with the well-supported sister group relations of *T. brunnea* and *T. montereyi* for mtDNA data (Hellberg 1998). The best Bayesian and ML trees for the VERL domain also groups *T. brunnea* + *T. montereyi* as sisters, but without significant support. For the Japanese trio of species, both the ZP domain and full VERL datasets support the sister group relationship of *T. lischkei* and *T. xanthostigma*, in agreement with mtDNA data. Bayesian analysis of the VERL domain supports an alternative topology, with *T. rusticus* the sister to *T. xanthostigma*, but support for this is marginal (BPP = 0.7, ML bootstrap < 0.7), and the ML analysis of the same data support a third alternative (*T. rusticus* + *T. lischkei*, not shown in Fig. 3). The relatively low resolution of the VERL domain does not appear to be related to the number of phylogenetically informative sites for each of the gene regions, as its value (109 sites) was intermediate to that for lysin (210) and the ZP domain (58).

INTERSPECIFIC PROTEIN DIVERGENCE AND POSITIVE SELECTION

Interspecific divergence of the protein sequence implicated in species-specific recognition (lysin and the N-terminal end of the putative VERL) was far greater than for the ZP domain of VERL or even for divergence of mitochondrial COI (Fig. 4). The rapid initial divergence of lysin and the VERL domain were especially evident among species within each of the three clades from which we had sampled three species (Table 2). In the extreme, the *T. brunnea*/*T. montereyi* species pair, less than 5% divergent at COI, had amino acid sequences that were almost 25% divergent for the VERL domain and nearly 50% divergent for the lysin protein (Table 2).

In contrast to lysin and the N-terminus of VERL, amino acid sequences in the ZP domain of VERL, which has not been implicated in any recognition function, diverge at a rate slower than that for mitochondrial COI (Fig. 4; Table 2).

The middle section of short repeats within VERL could only be compared for *T. funebralis* and *T. rusticus*; repeated efforts to amplify and sequence through this portion of the cDNA failed for the other species. As in *T. funebralis*, in *T. rusticus* this region is composed of short repeats, rich in serines and threonines, but these number only 116 residues in total (as compared to 244 in *T. funebralis*) and, along with the SPTT motif seen in *T. funebralis*, SATTSPTP and IPTT were common repeated motifs.

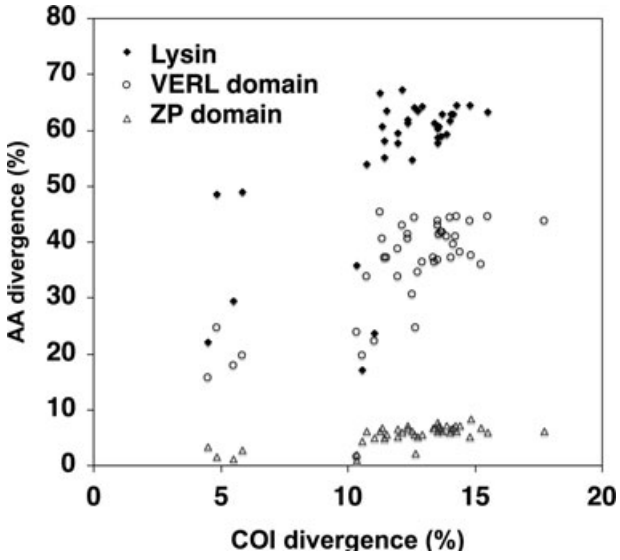

Figure 4. Protein divergence of lysin and two regions of VERL plotted against nucleotide divergence at mitochondrial COI. Amino acid divergence for the sperm protein lysin and the VERL domain that it interacts with are faster than for the mitochondrial gene. The ZP domain of VERL divergences more slowly than the mitochondrial gene.

Consistent with the extensive divergence of protein sequences for lysin and the VERL domain, sequence comparisons suggest that positive selection promotes the divergence of these regions. Pairwise comparisons of the ratio of nonsynonymous substitution per nonsynonymous site to synonymous substitutions per synonymous site (d*N*/d*S*, or ω) within the three clades (Table 2) are significantly greater than one for many species pairs. The d*N*/d*S* for lysin cDNAs was greatest among closely related Japanese species (2.64 between *T. lischkei* and *T. rusticus*, 3.10 between *T. lischkei* and *T. xanthostigma*), whereas those for the VERL domain were all very high (>5.8) for the three species in the subtidal Californian clade (*T. brunnea*, *T. montereyi*, and *T. regina*).

Likelihood analyses confirm that lysin and the VERL domain are under strong positive selection (Table 3). Overall, based on model M0, d*N*/d*S* was greater than one for lysin (1.66, 95% CIs 2.344–1.410) and for the VERL domain (1.81, 95% CIs 1.079–0.597), but well below one (0.25) for the ZP domain of VERL. The M8 model, which includes rate classes with d*N*/d*S* less than, equal to, and greater than one, estimated that nearly half the residues in lysin (47.6%) are evolving under positive selection, and that average d*N*/d*S* for those residues was 4.36. A smaller proportion of the VERL domain was inferred to be evolving via positive selection (26.3%), but d*N*/d*S* for these residues was even higher (5.70). Comparing model M7 (neutral and negatively selected rate classes) to model M8 (with an additional class for positively selected sites) suggests that positive selection is significant

Table 2. Pairwise amino acid divergence and ω (= d*N*/d*S*) estimates for *Tegula* lysin and the N-(containing the VERL domain) and C-(containing the ZP domain) termini of VERL.

		Lysin		VERL domain		ZP domain	
Species comparison	COI dist. (%)	*I* (%)	ω	*I* (%)	ω	*I* (%)	ω
bru-mon (S)	4.83	51.5	1.18	75.4	7.29**	98.5	0.22
bru-reg (A)	5.86	51.1	1.78**	80.3	5.51**	97.2	0.40
mon-reg (A)	5.51	70.6	1.57	82.0	46.0**	98.8	0.31
rus-lis (S)	11.04	76.4	2.74**	77.7	1.63	95.1	0.28
rus-xan (S)	10.55	82.9	1.39	80.2	1.31	95.7	0.22
lis-xan (S)	4.50	77.9	3.23**	84.3	4.42**	96.6	0.26
fun-gal (S)	10.33	64.2	1.52*	76.2	0.87	98.1	0.39
fun-rug (A)	12.67			75.4	0.75	97.8	0.22
gal-rug (A)	10.35			98.4	0.14	99.1	0.16

A = species pair with fully allopatric geographic distributions;
S = species pair with at least partially sympatric geographic distributions;
***Significant at the *P* < 0.05 level;**
****Significant at the *P* < 0.005 level.**

for lysin and the VERL domain, but not for the ZP domain (Table 3def).

A smaller number of residues met the more stringent criterion of a 95% posterior probability of evolving under positive selection: 42 residues in lysin (Fig. 5) and 20 residues from the VERL domain (Fig. 6). Inspecting these alignments reveals that different residues under selection change among close species in different clades. For example, residues 97–103 in lysin (Fig. 5) are different at nearly every residue in the subtidal Californian species (*T. brunnea*, *T. montereyi*, and *T. regina*), but almost perfectly invariant among the three Japanese species (*T. rusticus*, *T. lischkei*, and *T. xanthostigma*). Similar interclade differences occur for the VERL domain (Fig. 6): residues 71–76 are highly variable for the Japanese clade, but contain just a single replacement for the subtidal Californian clade. Such changes at different points in the primary sequence need not imply changes to different regions of proteins, depending on how the proteins are folded.

Table 3. Tests for positive selection on *Tegula* lysin and on the N-(VERL domain) and C-(ZP domain) termini of the putative *Tegula* VERL.

Gene/region	*L*	d*N*/d*S*	2Δ*L* (M8vM7)	2Δ*L* (M8vM8a)	P_1	ω
Lysin	142	1.66	93.8*	88.3*	0.476	4.36
VERL domain	122	1.81	81.2*	79.6*	0.263	5.70
ZP domain	324	0.25	1.1	0.6	0.103	1.52

***Significant at *P* = 0.005;**
***L* = number of codons;**
P_1 = proportion of sites under selection (under model M8);
ω = ratio of d*N*/d*S* for those sites under selection (model M8).

Overall, the d*N*/d*S* ratio of lysin and VERL in different species are correlated. The slope of the free model (about 8) was positive and significant. Pairwise values of d*N*/d*S* (Table 2) suggest selection may differ among the three clades. For example, although all pairwise comparisons among the three species in the Japanese and Californian subtidal clades are significantly greater than 1, d*N*/d*S* is below 1 for all of the comparisons of the VERL domain for the intertidal Californian clade. However, a three-clade model for different rates is not significantly better than the one-clade free model ($\chi^2 = 7.4$, df = 3).

Plotting inferred replacements to lysin and the VERL domain on the same phylogenetic tree (Fig. 7) suggests variation not only in the degree of positive selection, but also in the relative number of changes to the two interacting proteins. Among the lineages of the subtidal California clade, changes to lysin appear about three times as numerous as those to the VERL domain. Replacements in lysin and the VERL domain appear more nearly equal among the Japanese species. Few replacements distinguish the VERL domains of the (allopatric) sister species *T. gallina* and *T. rugosa*, even though they are twice as divergent at COI as the two other (sympatric) pairs of sister taxa.

Discussion

The two genes examined here are among both the many reproductive genes that show rapid interspecific divergence promoted by natural selection (Clark et al. 2006) and a more exclusive set for which interacting male and female genes have been examined. The *Tegula* sperm protein lysin is orthologous to *Haliotis* (abalone) lysin based on shared function and threading of the *Tegula* amino acid sequence on to the known crystallographic

```
                 1          11         21         31         41         51         61
                 | *  ** *  *     **   |   *   *  |  * **    |  ** *    |          |*
T. aureotincta   SMR-VPIVRH GNVDFGRSEN GWIKRGAVEE MDKQADKYVR ERPNLRSYIP MFKYFSKMKV YNMWPNWSNW
T. brunnea       --HR-ISNVN S.K...GVN. .VM.TAI.KA LH.K.TVWC. QH.HG.P.E. FMRFMNVQR. .TN.N.M.T.
T. montereyi     NGYRPVFKAR TTK....AN. .I..TYIIK. ..HK.HM.I. KH.EG.P.L. FMRFMN.Q.L .MT.N...T.
T. regina        CGSRPQVTVL S.K....TN. .I..T.IIK. ..HK.AM... KH.SA.P.E. FM..MN.Q.L .TT.N...T.
T. rusticus      YGPG.RV.V. ..K...YKN. .I..SAI.RT L.RFS.A... KH.ST.K.KD YMMFLNRR.. ..T.N...S.
T. lischkei      YGPG.KV.V. Q.K.L.QKN. .IV.EAI.RT ..RFV.A... KH.SS.R.KD YMR.MNRR.. ..T.N.....
T. xanthostigma  YGAG.RV.VA N.R...YEN. .IV.TAI.RV ..RFS.A... KH.SA.R.LD YMRFLNRR.. ..T.N.....
T. funebralis    HTPG.R..Q. .Y.....VN. .L..G.MFL. ..NV.KRFC. KH.SAKP.FQ YMR.LNRQRI .GN.N.Y.Q.
T. gallina       HTPG.VV... .Y.....HN. .L..A.MFH. ..HK.RV.C. S..KA.P.LD YMR.IHRRMI .LN.N.Y.Q.
T. rugosa        HTPG.VV... .Y.....HN.

                71         81         91         101        111        121        131        141
                |**     *  | *    *   | *   **** *** **  ** ****       |     * ** |        * |*
T. aureotincta  CATWLRRLNR TPHARDYAAC GKIRGRE-AY MPHLYDVAVR QNYKTLNPYE KKILATAPIH L-PIRAV-GR FA
T. brunnea      AVKE..KMH. R.VT...ENL .RRI.-.HT. .R.V.E.VSE MRIR-PT.DQ IRFTNIK.AN ...L.T-P.. .G
T. montereyi    .SKEIAKM.. RAT....ENL .RRL.KIVYM DLAYRV.VQL RMQ--P..DQ RRF.NLRAAD .....TFP.H --
T. regina       ..KEIAKMG. K.NS...ENL .RRL.KLAYM DFAYSV..RL RMQ--...AQ RRF.YIK.AD ...V.T...K YF
T. rusticus     ..HRI.EM.. K.TS..FDSF .VRL.NMCRF QMNFMYDVIV KERCK..AHQ R.F.N.P.PD M....-.P.. .F
T. lischkei     .INRI.E... R.N.G.FRRF .VRLANMCR. QMNFMYDV.V KERSQ..ANQ RNF.N.P.AR M....-.P.. ..
T. xanthostigma .RNRI.E... R.T...FRSF .ERL.NMCRF QMNFMYDVIV KERCK..AHQ RRF.N.P.RD M....-.P.K .F
T. funebralis   ARGLVQK.G. K.TS.EF.NI .RKM.K.MDC EA-YFRIV.. Y-RLK...DK R.L.N.-.AI DF...TL.RN WG
T. gallina      ARMLVQK.R. K.SS...TNI ..VM.KKIYT NGYYPEIVKL --RCKF..DK RNF.N.-.AV DL.L.TLRK. WG
```

Figure 5. Alignment of the mature lysin proteins from ten species of *Tegula*. Residues marked with an asterisk are indicated as being under positive selection by a BEB analysis with a posterior probability of 0.95 or greater.

```
                1          11         21         31         41         51
                |          *  *      |          | *   *    | * * *    |*
T. aureotincta  AKNEQSLVSL EGINVTDVLA LGNRFMALNI SKNCSQDIDL PSRFELKFDY SKLNVTVFNV
T. brunnea      .E.......M K..SMI.I.. ....VVG... T.E..DNFGE ..S.K.V... TL.....V..
T. montereyi    .E....F... N..SLN.I.. ....VVG... T.K..DT..E ..S.S.L... TL.....L..
T. regina       .E.......M D.VSL..I.. ....VVG... T.Q..DTF.E ..T...L... TR....AL..
T. rusticus     .......... D.V......D ....VFGV.. T.K..DTFGD ..S.QFV... .L.....L.I
T. lischkei     .........M D.VP..EI.. ..D.LVG... T.K..DNF.D ..S.QFV..N .L.....L.I
T. xanthostigma .........M D.V...QI.D ....VVG... T.Q..DTF.D ..S.QFV... .L.....L.I
T. funebralis   .V........ K.VS.RE... ....INGM.. T....D.F.E L.E.NFY... TV.....V.I
T. gallina      .E........ K.VS.SEI.. ....IDG... T.D..DNF.V S.K.YFN... TR.....V.I
T. rugosa       .E........ K..S.SEI.. ....IDG... T.D..DNF.V S.K.YFN... TR.....V.I

                61         71         81         91         101        111        121
                |       *  *****  * * |  *       | * *      |    *     |          |
T. aureotincta  SSACLGSTGI TDRIGN-TFI GSVFHPDGKS DVVSKCYFTP VVTDTTMKNI VYFLVEWADG KV
T. brunnea      T.S...D.VL PQNVNDKFYL A..Y..EES. ..T.D...S. ..S.AAL.H. L..QI..S.. .I
T. montereyi    T.S..DT.MS PQKVNDKMYV A..Y..EES. .TE.....  .LS..VL.H. MI.....S.. .I
T. regina       T.S...T.VL PQKVNDKLYL A..Y..EES. .IE.N.....  ..S...L.H. L......S.. .I
T. rusticus     ........QL LQLTQD.SYT AF.S...ENV NYK.Q.....  ..S...L.H. .F..L..S.. .I
T. lischkei     ........QL AEKSAS.SYT AF.....EGI ..I.Q.....  ..S..IL.HV ....L..S.. .I
T. xanthostigma ........QL AQKS.E.SYT AF.Y...QSI ..S.Q.F...  ..S.RIL.H. ....L..S.. ..
T. funebralis   .L......FD SM.VEDDH.A .LAY..EAS. P.EN....A. I.S.MQV.H. L...I..S.. .L
T. gallina      .L......FD SVKLDDDA.M .F.L..EAS. P.STT.....  ..S.MKL.H. L...I..S.. .L
T. rugosa       .L......FD SVKLDDDA.M .F.L..EAS. P.STS.....  ..S.MKL.H. L...I..S.. .L
```

Figure 6. Alignment of the amino terminus of the putative VERL protein from ten species of *Tegula*. Residues marked with an asterisk are indicated as being under positive selection by a BEB analysis with a posterior probability of 0.95 or greater.

structure of *Haliotis* lysin (Hellberg and Vacquier 1999). The *Tegula* VE protein isolated here appears to be orthologous to that from *Haliotis* based on shared possession of a unique amino acid motif (previously seen only in two *Haliotis* paralogs that both bind to lysin, Aagaard et al. 2010) and phylogenetic analysis of a ZP domain shared by many gastropod egg-coat proteins (Aagaard et al. 2010). Like the lysin/VERL pair from abalone, the interacting proteins in *Tegula* both diverge very rapidly between species (with amino acid replacement rates an order of magnitude greater than mtDNA substitution rates in some cases, Fig. 4) and levels of positive selection acting on each are correlated. Although the mechanism responsible for this positive selection remains unclear, aspects of the data are consistent with roles for both reinforcement and sexual anatagonism.

HOMOLOGY OF *TEGULA* AND *HALIOTIS* VERL

The homology of the *Tegula* protein characterized here and *Haliotis* VERL is supported by shared patterns of protein abundance and biochemical properties, phylogenetic analysis of conserved sequence, and a shared unique motif. Like the *Haliotis* protein,

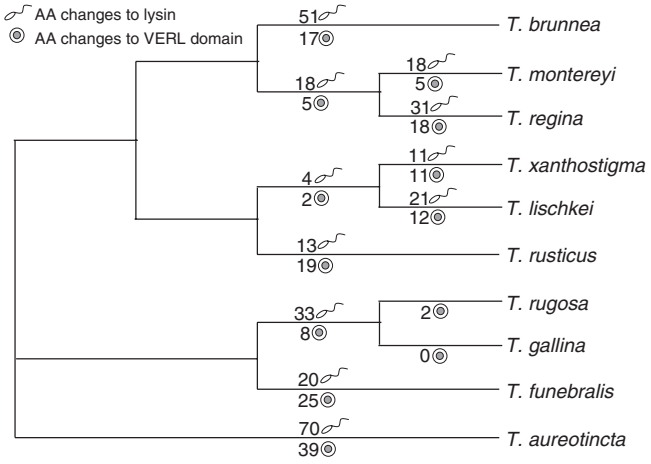

Figure 7. Inferred number of amino acid replacements to lysin (above branch) and the VERL domain (below branch) along a phylogenetic tree of the *Tegula* species inferred from the genes encoding lysin and VERL.

Tegula VERL is the most abundant protein in the VE, the layer outside the egg proper that an approaching sperm would first encounter. *Tegula* VERL is also very large, with a PAGE-estimated size >1 MDa. This size is far larger than the inferred molecular weight of 77.6 KDa calculated from the cDNA sequence of *T. funebralis*. *Haliotis* VERL likewise shows a large discrepancy between the PAGE-estimated size of the protein and that inferred from cDNA sequence (Galindo et al. 2002), a difference attributed to high levels of O-glycosylation. Both VERLs have a large number of serine and threonine residues that could be glycosylated. The role carbohydrates may play in lysin-VERL binding is not known, but tryptic glycopeptides of *Haliotis* VERL do not inhibit lysin binding (W. J. Swanson, unpubl. data from *Haliotis*), suggesting that it is the peptide backbone of VERL, not merely the carbohydrates bound to it, that is critical to its interactions with lysin.

The *Tegula* VERL protein consists of three major regions: a unique amino-terminus of about 115 residues, a central region composed of many repeats of a three or four residues, and a carboxy-terminal ZP domain. Short motifs have been implicated in the interactions of sperm-egg proteins previously (Vacquier et al. 1996; Palumbi 1999) and insertions and deletions of such motifs have been shown to be under positive selection in other sex proteins (Dixon Schully and Hellberg 2006). Short motifs rich in serine, proline, and threonine are typical of O-glycosylated proteins (Perez-Vilar and Hill 1999), although sharing such residues in repeats is more likely to indicate common function than shared ancestry because they have likely evolved many times independently. Repeats rich in S, P, and T form the core of copulatory plug-forming mucoproteins in nematodes (Palopoli et al. 2008) that, like VERL, may bind to sperm. A paralog of VERL in *Haliotis*, VEZP14 (Aagaard et al. 2010), possesses a mid-section of about 20 repeats of TTTTTTP. The divergence of repeats seen here between *T. funebralis* and *T. rusticus* correlates with extensive sequence changes to lysin; however, the wide variation in both the length of repeat units (from three to eight residues within the two *Tegula* examined) and the total length of the mid-section (between 116 and 244 residues) stands in contrast to the conserved size of *Tegula* lysins (all between 137 and 140 residues long).

The two ends of *Tegula* VERL provide firmer support for its homology to *Haliotis* VERL. The ZP domain is common to many egg-coat proteins, not only in molluscs (Aagaard et al. 2006; Sedik et al. 2010), but also in tunicates (Yamada et al. 2009) as well as in mammals, from which they were originally described (Monne et al. 2006) and in mating proteins from fungi (Swanson et al. 2011). Such strong conservation suggests that the ZP domain plays a structural role in forming the egg coat proper (Swanson et al. 2011), a role likely to impose constraints on the evolution of this portion of the VERL protein. Phylogenetic analysis of this ZP domain from *Tegula* VERL, *Haliotis* VERL, two *Tegula* VE proteins (Haino-Fukushima et al. 2000), and 32 additional *Haliotis* VE proteins placed *Tegula* VERL as the sister to *Haliotis* VERL and VEZ14 (Aagaard et al. 2010). *Tegula* VERL and the *Haliotis* VERL and VEZP14 duo are also the only proteins known to share the domain at their amino-terminal ends.

In *Haliotis*, lysin and another abundant and positively selected 18-KDa acrosomal protein appear to be recent duplicates (Swanson and Vacquier 1995). The egg-borne proteins VERL and VEZ14 may have duplicated in tandem to match these sperm proteins (Aagaard et al. 2010). Given the closeness of VERL and VEZ14 relative to *Tegula* VERL, such a scenario would imply that *Tegula* VERL serves either the same function as one of the two *Haliotis* proteins or a mosaic of both. VERL and VEZ14 are the *Haliotis* VE proteins that bind tightest to lysin and 18-KDa affinity columns (Aagaard et al. 2010), so the *Tegula* VERL described here is likely to interact with lysin, a conclusion bolstered by the large fraction of the VE this protein composes. Recent duplications in the lineage leading to *Haliotis* would also imply that *Tegula* would posses a single protein homologous to lysin, rather than the two seen in *Haliotis*. Although *Tegula* sperm proteins have not been surveyed exhaustively, the major acrosomal protein in *Tegula* (besides lysin) is not homologous to lysin (Hellberg et al. 2000).

Together, these observations suggest that the putative VERL from *Tegula* described here is orthologous to that from *Haliotis*, and that the unique amino-terminus they share is the region of the protein that determines the specificity of binding to lysin. Direct demonstration that this is so will require biochemical analysis of the specificity of binding of different regions of VERL to homo- and heterospecific lysins.

POSITIVE SELECTION ON LYSIN AND VERL

Proteins functionally tied to sex and reproduction are frequently targets of positive selection (Swanson and Vacquier 2002). Most early examples were male-specific proteins, but the signal of positive selection has also become evident in female-specific proteins (Swanson et al. 2001; Swanson et al. 2004; Turner and Hoekstra 2006). Although coevolution between male and female reproductive traits is well known (e.g. Pitnick et al. 2003 for *Drosophila*; Beese et al. 2006 for copulating pulmonate snails), few studies have ascribed selected nucleotide substitutions in one sex to changes in the other. Briscoe et al. (2010) provide a recent example, in which selection promotes divergence on a duplicated ultraviolet-sensitive opsin gene from *Heliconius* butterflies that appears to have arisen coincident with the origin of novel wing pigments. These butterflies are known to use visual cues to identify mates, so these parallel changes may stem from the selective pressures of courtship. More direct comparisons between interacting male and female reproductive proteins are rarer still, with the sole metazoan example coming from the lysin and VERL of abalone (Clark et al. 2009).

Levels of positive selection (d*N*/d*S*) along branches of a shared species topology are correlated for *Tegula* lysin and the rapidly evolving amino-terminus of VERL. The slope of this relationship (near 8) is similar to that seen between lysin and VERL in abalone (around 10; Clark et al. 2009) and is consistent with the observation that a higher proportion of male sex proteins show high d*N*/d*S* ratios compared to female sex proteins (Swanson et al. 2001; 2004). Civetta (2003) suggested that the relative rates of interacting sex proteins could be used to judge whether the mechanisms underlying their divergence is female-driven sexual selection or male-driven sexual antagonism. Both Civetta (2003) and Clark et al. (2009) recognized, however, that such straightforward comparisons were complicated by the repetition of the putative lysin-binding motif in *Haliotis* VERL. *Tegula*, with just a single such motif, offer a more straightforward comparisons, especially given the similar sizes of lysin and the amino-terminal portion of VERL with which it interacts. Such comparisons suggest a faster rate of male evolution here, consistent with sexual antagonism, although the hand of other selective forces cannot be discounted and caution is warranted until a fuller understanding of the stoichiometry of lysin-VERL binding has been obtained in *Tegula*.

An independent line of reasoning lends support to sexual antagonism underlying at least some of the rapid rates of interspecific divergence in *Tegula* lysin and VERL. Sexual conflict can arise within externally fertilized species when males compete for eggs and eggs are at risk from polyspermy (Levitan et al. 2007; Levitan 2008). In species that spawn at high gamete concentrations, intense sperm competition will select for fast fertilization by males that will be in conflict with selection against polyspermy in females. At low gamete densities, where fertilization success is low for both sexes, both sexes should be under sexual selection. We would thus expect rates of changes to male sex proteins to be greater than those of females at low gamete concentrations, with female rates being higher at high gamete concentrations when selection favors stronger choosiness.

These predictions are in accord with variation in the relative numbers of inferred amino acid substitutions among different lineages within *Tegula* (Fig. 7). In the *brunnea*/*montereyi*/*regina* clade, where all species occur subtidally and thus might be expected to encounter low gamete concentrations, lysin shows two or three times as many amino acid replacements as the VERL domain. In contrast, in the intertidal *funebralis*/*gallina*/*rugosa* and *rusticus*/*xanthostigma*/*lischkei* clades, where spawning at low tides in pools could lead to high gamete concentrations, relative numbers of changes are more comparable, with replacements in VERL exceeding those to lysin in some lineages. Similar qualitative patterns in relative differences remain when only selected residues are considered (not shown). These differences in relative rates of change are correlated not only with environment, but measures of adult density, which can be over 10 times greater for the high intertidal *T. funebralis* (>400 per m^2; Frank 1975) than for subtidal species such as *T. brunnea* and *T. montereyi* (Watanabe 1982), with low intertidal species falling between these (Schmitt 1996). Although consistent with the predictions and experimental results of Levitan et al. (2007), the conclusion drawn here is necessarily weak, given the single comparison and the large intraspecific variation in the densities of breeding adults within some *Tegula* species (e.g., *T. funebralis*, Cooper and Shanks 2011).

A fuller understanding of the mechanisms driving the divergence of sex proteins and their possible role in the evolution of reproductive isolation may come from study of the phylogenies of interacting genes. Genes functionally tied to reproductive isolation should prove reliable markers for species identity if they perform poorly in a heterospecific genetic background and selection within species reduces variation in a way that limits lineage sorting. Such genes may be marked by limited introgression across hybrid zones (Rieseberg et al. 1999; Carling and Brumfield 2009; Maroja et al. 2009) and strong allele sorting in regions of sympatry with close relatives (Geyer and Palumbi 2003), and their gene trees may trace the true species tree more closely than those of neutral markers (Ting et al. 2000; Palopoli et al. 1996).

Here, some of the phylogenetic trees based on sequences for lysin or VERL (Fig. 3) differ from previously published topologies based on mitochondrial gene sequences (see Fig. 3 in Hellberg 1998, in section of *Gene trees*). Most striking is the change in species falling as sister to *T. montereyi*. Mitochondrial trees (Hellberg 1998) based on several reconstruction methods found ≥95% bootstrap support for a pairing with *T. brunnea*. Only the VERL domain recovers such a relationship here, and

with weak support, whereas a sister pairing with *T. regina* receives rather strong support from both lysin and the ZP domain of VERL (Fig. 3). At a simple gene marker level, this conflict exemplifies the common problem of investing too much faith in a single gene tree for revealing species relationships, and may solve the previously stated mystery (Hellberg 1998) of how the geographic range of one species (*T. montereyi*) could be completed nested within that of its sister (supposedly *T. brunnea*); *T. regina* and *T. montereyi* are allopatric.

Closer to a mechanistic understanding of how gamete recognition proteins diverge during speciation, the varying degrees of resolution and conflict revealed by the different gene trees may provide hints about levels of variation carried at the interacting loci during divergence. Intraspecifc variation in abalone is very low (Metz et al. 1998b; Clark et al. 2009). There are no published data on lysin variation in *Tegula*, although a preliminary survey in *T. funebralis* from Oregon and Baja California likewise found fixation at the amino acid level (M. E. Hellberg, unpubl. data) and little variation among nucleotide sequences. In the extreme, strong purifying selection that creates fixation within species should negate the possibility of lineage sorting at a locus, although the forces that normally enforce fixation within species may become disrupted during the divergence process.

Patterns of gene tree resolution, protein divergence, and positive selection do not appear to be independent of one another. For example, the ZP domain of VERL, which does not experience positive selection, shows greater gene tree resolution at all nodes than does the VERL domain, which shows strong positive selection (Fig. 3; Table 2). Such is also the case for the two selected repeats in *Haliotis* VERL (see Fig. 1 in Galindo et al. 2003). Does this mean that VERL carries more variation through the speciation event than the ZP domain in the same gene, or than the male's lysin with which it interacts? The *brunnea/montereyi/regina* clade that most often shows poor nodal support and conflict among gene trees (Fig. 3) also shows the highest values of and the greatest rates of VERL and lysin divergence relative to mtDNA divergence (Table 2). Does this suggest continuing selection for divergence during an episode of secondary contact and mitochondrial introgression? Resolving these questions will require resolution of the *Tegula* species tree and a theoretical exploration of expected patterns of gene tree coalescence under different models of protein interaction, population isolation, and demographic history, but it seems possible that the same lysin/VERL recognition system that first revealed a strong signature of positive selection may one day also help expose the mechanisms by which interacting sex proteins diverge during the evolution of reproductive isolation.

ACKNOWLEDGMENTS

This work was supported by DEB-0075382 to MEH, OCE-0550270 to MEH and I. Baums, NIH grant HD057974 and HD42563 to WJS. JEA was supported by NSF grant DEB 0918106. We thank R. Eytan, C. Prada, S. Springer, J. Larkin, two anonymous reviewers, and especially V. Vacquier for helpful discussions and criticisms.

LITERATURE CITED

Aagaard, J. E., X. Yi, M. J. MacCoss, and W. J. Swanson. 2006. Rapidly evolving zona pellucida proteins are a major component of the vitelline envelope of abalone eggs. Proc. Natl. Acad. Sci. USA 103:17302–17307.

Aagaard, J. E., V. D. Vacquier, M. J. MacCoss, and W. J. Swanson. 2010. ZP domain proteins in the abalone egg coat include a paralog of VERL under positive selection that binds lysin and 18-kDa sperm proteins. Mol. Biol. Evol. 27:193–203.

Aguadé, M. 1999. Positive selection drives the evolution of the Acp29AB accessory gland protein in *Drosophila*. Genetics 152:543–551.

Beese, K., K. Beier, and B. Baur. 2006. Coevolution of male and female reproductive traits in a simultaneously hermaphroditic land snail. J. Evol. Biol. 19:410–418.

Briscoe, A. D., S. M. Bybee, G. D. Bernard, F. Yuan, M. P. Sison-Mangus, R. D. Reed, A. D. Warren, J. Llorente-Bousquets, and C.-C. Chiao. 2010. Positive selection of a duplicated UV-sensitive visual pigment coincides with wing pigment evolution in *Heliconius* butterflies. Proc. Natl. Acad. Sci. USA 107:3628–3633.

Carling, M. D., and R. T. Brumfield. 2009. Speciation in *Passerina* buntings: introgression patterns of sex-linked loci identify a candidate gene region for reproductive isolation. Mol. Ecol. 18:834–847.

Chomczynski, P., and N. Sacchi. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. 162:156–159.

Civetta, A. 2003. Shall we dance or shall we fight? Using DNA sequence data to untangle controversies surrounding sexual selection. Genome 46:925–929.

Clark, N. L., and W. J. Swanson. 2005. Pervasive adaptive evolution in primate seminal proteins. PLoS Genet. 1:335–342.

Clark, A. G., D. J. Begun, and T. Prout. 1999. Female × male interactions in *Drosophila* sperm competition. Science 283:217–220.

Clark, N. L., J. E. Aagaard, and W. J. Swanson. 2006. Evolution of reproductive proteins from animals and plants. Reproduction 131:11–22.

Clark N. L., J. Gaspar, M. Sekino, S. A. Springer, C. F. Aquadro, and W. J. Swanson. 2009. Coevolution of interacting fertilization proteins. PLoS Genet. 5:e1000570.

Cooper, E. E., and A. L. Shanks. 2011. Latitude and coastline shape correlate with age-structure of *Chlorostoma* (*Tegula*) *funebralis* populations. Mar. Ecol. Prog. Ser. 424:133–143.

Coyne, J. A., and H. A. Orr. 2004. Speciation. Sinauer: Sunderland, MA. 543 p.

Dixon Schully, S., and M. E. Hellberg. 2006. Positive selection on nucleotide substitutions and indels in accessory gland proteins of the *Drosophila pseudoobscura* subgroup. J. Mol. Evol. 62:793–802.

Folmer, O., M. Black, W. Hoeh, R. Lutz, and R. Vrijenhoek. 1994. DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. Mol. Mar. Biol. Biotechnol. 3:294–299.

Frank, P. W. 1975. Latitudival variation in the life history features of the black turban snail *Tegula funebralis* (Prosobranchia: Trochidae). Mar. Biol. 31:181–192.

Galindo, B. E., G. W. Moy, W. J. Swanson, and V. D. Vacquier. 2002. Full-length sequence of VERL, the egg vitelline envelope receptor for abalone sperm lysin. Gene 288:111–117.

Galindo, B. E., V. D. Vacquier, and W. J. Swanson. 2003. Positive selection in the egg receptor for abalone sperm lysin. Proc. Natl. Acad. Sci. USA 100:4639–4643.

Geyer, L. B., and H. A. Lessios. 2009. Lack of character displacement in the male recognition molecule, bindin, in Atlantic sea urchins of the genus *Echinometra*. Mol. Biol. Evol. 26:2135–2146.

Geyer, L. B., and S. R. Palumbi. 2003. Reproductive character displacement and the genetics of gamete recognition in tropical sea urchins. Evolution 57:1049–1060.

Goldman, N., and Z. Yang. 1994 A codon-based model of nucleotide substitution for protein-coding DNA sequences. Mol. Biol. Evol. 11:725–736.

Haino, K. 1971. Studies on the egg-membrane lysin of *Tegula pfeifferi*. Purification and properties of the egg-membrane lysin. Biochim. Biophys. Acta 229:459–470.

Haino-Fukushima, K. 1974. Studies on the egg-membrane lysin of *Tegula pfeifferi*: the reaction mechanism of the egg-membrane lysin. Biochim. Biophys. Acta 352:179–191.

Haino-Fukushima, K., M. Shimoirisa, and Y. Yamakawa. 1999. Vitelline-coat lysin: comparisons of lysins among three species in the genus *Tegula*. Dev. Growth Differ. 41:455–461.

Haino-Fukushima, K., X. Fan, and S. Nakamura. 2000. Three new components contained in the vitelline coat of *Tegula pfeifferi*. Zygote 8:S61.

Hellberg, M. E. 1998. Sympatric sea shells along the sea's shore: the geography of speciation in the marine gastropod *Tegula*. Evolution 52:1311–1324.

Hellberg, M. E., and V. D. Vacquier. 1999. Rapid evolution of fertilization selectivity and lysin cDNA sequences in teguline gastropods. Mol. Biol. Evol. 16:839–848.

Hellberg, M. E., G. W. Moy, and V. D. Vacquier. 2000. Positive selection and propeptide repeats promote rapid interspecific divergence of a gastropod sperm protein. Mol. Biol. Evol. 17:458–466.

Howard, D. J. 1999. Conspecific sperm and pollen precedence and speciation. Annu. Rev. Ecol. Syst. 30:109–132.

Huelsenbeck, J. P., and B. Rannala. 2004. Frequentist properties of Bayesian posterior probabilities of phylogenetic trees under simple and complex substitution models. Syst. Biol. 53:904–913.

Jaffe, L. A., and M. Gould. 1985. Polyspermy preventing mechanisms. Pp. 223–250 *in* C. B. Metz and A. Monroy, eds. Biology of fertilization, Vol. 3. Academic Press, San Diego, CA.

Jagadeeshan, S., and R. S. Singh. 2007. Rapid evolution of outer egg membrane proteins in the *Drosophila melanogaster* subgroup: a case of ecologically driven evolution of female reproductive traits. Mol. Biol. Evol. 24:929–938.

Jovine, L., C. C. Darie, E. S. Litscher, and P. M Wasserman. 2005. Zona pellucida domain proteins. Annu. Rev. Biochem. 74:83–114.

Kamei, N., and C. G. Glabe. 2003. The species-specific egg receptor sea urchin sperm adhesion is EBR1, a novel ADAMTS protein. Genes Dev. 17:2502–2507.

Kelleher, E. S., W. J. Swanson, and T. A. Markow. 2007. Gene duplication and adaptive evolution of digestive proteases in *Drosophila arizonae* female reproductive tracts. PLoS Genet. 3:1541–1549.

Kresge, N., V. D. Vacquier, and C. D. Stout. 2000. 1.35 and 2.07 Å resolution structures of the red abalone sperm lysin monomer and dimer reveal features involved in receptor binding. Acta Crystallogr. D56:34–41.

———. 2001. Abalone lysin: the dissolving and evoluting sperm protein. BioEssays 23:95–103.

Lee, Y.-H., T. Ota and V. D. Vacquier. 1995. Positive selection is a general phenomenon in the evolution of abalone sperm lysin. Mol. Biol. Evol. 12:231–238.

Levitan, D. R. 2008. Gamete traits influence the variance in reproductive success, the intensity of sexual selection, and the outcome of sexual conflict among congeneric sea urchins. Evolution 62:1305–1316.

Levitan, D. R., and D. L. Ferrell. 2006. Selection on gamete recognition proteins depends on sex, density, and genotype frequency. Science 312:267–269.

Levitan, D. R., and A. P. Stapper. 2010. Simultaneous positive and negative frequency-dependent selection on sperm bindin, a gamete recognition protein in the sea urchin *Strongylocentrotus purpuratus*. Evolution 64:785–797.

Levitan, D. R., C. P. terHorst, and N. D. Fogarty. 2007. The risk of polyspermy in three congeneric sea urchins and its implications for gametic incompatibility and reproductive isolation. Evolution 61:2007–2014.

Lewis, C. A., C. F. Talbot, and V. D. Vacquier. 1982. A protein from abalone sperm dissolves the egg vitelline layer by a non-enzymatic mechanism. Dev. Biol. 92:227–239.

Maroja, L. S., J. A. Andres, and R. G. Harrison. 2009. Genealogical discordance and patterns of introgression and selection across a cricket hybrid zone. Evolution 63:2999–3015.

McCartney, M. A., and H. A. Lessios. 2004. Adaptive evolution of sperm bindin tracks egg incompatibility in Neotropical sea urchins of the genus *Echinometra*. Mol. Biol. Evol. 21:732–745.

Metz, E. C., and S. R. Palumbi. 1996. Positive selection and sequence rearrangements generate extensive polymorphism in the gamete recognition protein bindin. Mol. Biol. Evol. 13:397–406.

Metz, E. C., G. Gomez-Gutirrez, and V. D. Vacquier. 1998a. Mitochondrial DNA and bindin gene sequence evolution among allopatric species of the sea urchin genus *Arbacia*. Mol. Biol. Evol. 15:185–195.

Metz, E. C., R. Robles-Sikisaka, and V. D. Vacquier. 1998b. Nonsynonymous substitution in abalone sperm fertilization genes exceeds substitution in introns and mitochondrial DNA. Proc. Natl. Acad. Sci. USA 95:10676–10681.

Monne, M., L. Han, and L. Jovine. 2006. Tracking down the ZP domain: from the mammalian zona pellucida to the molluscan vitelline envelope. Sem. Repro. Med. 24:204–216.

Nei, M., and T. Gojobori. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol. Biol. Evol. 3:418–426.

Palopoli, M. F., A. W. Davis, and C.-I. Wu. 1996. Discord between the phylogenies inferred from molecular versus functional data: uneven rates of functional evolution or low levels of gene flow? Genetics 144:1321–1328.

Palopoli, M. F., M. V. Rockman, A. TinMaung, C. Ramsay, S. Curwen, A. Aduna, J. Laurita, and L. Kruglyak. 2008. Molecular basis of the copulatory plug polymorphism in *Caenorhabditis elegans*. Nature 454:1019–1022.

Palumbi, S. R. 1999. All males are not created equal: fertility differences depend on gamete recognition polymorphisms in sea urchins. Proc. Natl. Acad. Sci. USA 96:12632–12637.

Perez-Vilar, J., and R. L. Hill. 1999. The structure and assembly of secreted mucins. J. Biol. Chem. 274:31751–31754.

Pitnick, S., G. T. Miller, K. Schneider, and T. A. Markow. 2003. Ejaculate-female coevolution in *Drosophila mojavensis*. Proc. R. Soc. Lond. B 270:1507–1512.

Pond, S. L., S. D. Frost, and S. V. Muse. 2005. HyPhy: hypothesis testing using phylogenies. Bioinformatics 21:676–679.

Posada, D. 2008. jModelTest: phylogenetic model averaging. Mol. Biol. Evol. 25:1253–1256.

Posada, D., and T. R. Buckley. 2004. Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. Syst. Biol. 53:793–808.

Rambaut, A., and A. J. Drummond. 2010. Tracer v1.5. Available at http://beast.bio.ed.ac.uk/Tracer (accessed December 17, 2009).

Rice, W. R. 1996. Sexually antagonistic male adaptation triggered by experimental arrest of female evolution. Nature 381:232–234.

Rice, W. R., and B. Holland. 1997. The enemies within: intergenomic conflict, interlocus contest evolution (ICE), and the intraspecific Red Queen. Behav. Ecol. Sociobiol. 41:1–10.

Rieseberg, L. H, J. Whitton, and K. Gardner. 1999. Hybrid zones and the genetic architecture of a barrier to gene flow between two sunflower species. Genetics 152:713–727.

Riginos, C., and J. H. McDonald. 2003. Positive selection on an acrosomal sperm protein, M7 lysin, in three species of the mussel genus *Mytilus*. Mol. Biol. Evol. 20:200–207.

Ronquist, F., and J. P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574.

Schmitt, R. J. 1996. Exploitation competition in mobile grazers: trade-offs in use of a limited resource. Ecology 77:408–425.

Sedik, W. F., K. E Dempsey, X. Meng, and J. A. Craft. 2010. Temporal expression of sex-specific genes in the mantle of the common mussel (*Mytilus edulis*). Mar. Biol. 157:639–646.

Shaw, A., D. E. McRee, V. D. Vacquier, and C. D. Stout. 1993. The crystal structure of lysin, a fertilization protein. Science 262:1864–1867.

Springer, S. A., and B. J. Crespi. 2007. Adaptive gamete-recognition divergence in a hybridizing *Mytilus* population. Evolution 61:772–783.

Stekoll, M. S., and T. C. Shirley. 1993. In situ spawning behavior of an Alaskan population of pinto abalone, *Haliotis kamtschatkana* Jonas, 1845. Veliger 36:95–97.

Stephano, J. L. 1992. A study of polyspermy in abalone. Pp. 518–526 *in* S. A. Sheppard, M. J. Tegner, and S. A. Guzman, eds. Abalone of the world: biology, fisheries and culture. Blackwell, Cambridge, UK.

Swanson, W. J., and V. D. Vacquier. 1995. Extraordinary divergence and positive Darwinian selection in a fusagenic protein coating the acrosomal process of abalone spermatozoa. Proc. Natl. Acad. Sci. USA 92:4957–4961.

———. 1997. The abalone egg vitelline envelope receptor for sperm lysin is a giant multivalent molecule. Proc. Natl. Acad. Sci. USA 94:6724–6729.

———. 1998. Concerted evolution in an egg receptor for a rapidly evolving abalone sperm protein. Science 281:710–712.

———. 2002. Reproductive protein evolution. Annu. Rev. Ecol. Syst. 33:161–179.

Swanson, W. J., Z. Yang, M. F. Wolfner, and C. F. Aquadro. 2001. Positive Darwinian selection drives the evolution of several female reproductive proteins in mammals. Proc. Natl. Acad. Sci. USA 98:2509–2514.

Swanson, W. J., R. Nielsen, and Q. F. Yang. 2003. Pervasive adaptive evolution in mammalian fertilization proteins. Mol. Biol. Evol. 20:18–20.

Swanson, W. J., A. Wong, M. F. Wolfner, and C. F. Aquadro. 2004. Evolutionary expressed sequence tag analysis of *Drosophila* female reproductive tracts identifies genes subjected to positive selection. Genetics 168:1457–1465.

Swanson, W. J., J. E. Aagaard, V. D. Vacquier, M. Monné, H. S. Al Hosseini, and L. Jovine. 2011. The molecular basis of sex: linking yeast to human. Mol. Biol. Evol. 28:1963–1966.

Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei, and S. Kumar. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28:2731–2739.

Ting, C.-T., S.-C. Tsaur, and C.-I. Wu. 2000. The phylogeny of closely related species as revealed by the genealogy of a speciation gene, Odysseus. Proc. Natl. Acad. Sci. USA 97:5313–5316.

Tracey S., J. A. Todd, and D. H. Erwin. 1993. Mollusca: gastropoda. Pp. 131–167 *in* M. J. Benton, ed. The fossil record 2. Chapman and Hall, London.

Tsaur, S.C., and C.-I. Wu. 1997. Positive selection and the molecular evolution of a gene of male reproduction, *Acp26Aa* of *Drosophila*. Mol. Biol. Evol. 14:544–549.

Turner, L. M., and H. E. Hoekstra. 2006. Adaptive evolution of fertilization proteins within a genus: variation in ZP2 and ZP3 in deer mice (*Peromyscus*). Mol. Biol. Evol. 23:1656–1669.

———. 2008. Reproductive protein evolution within and between species: maintenance of divergent ZP3 alleles in *Peromyscus*. Mol. Ecol. 17:2616–2628.

Vacquier, V. D., K. R. Carner, and C. D. Stout. 1990. Species-specific sequences of abalone lysin, the sperm protein that creates a hole in the egg envelope. Proc. Natl. Acad. Sci. USA 87:5792–5796.

Vacquier, V. D., W. J. Swanson, and M. E. Hellberg. 1996. What have we learned about sea urchin sperm bindin? Dev. Growth & Differ. 37:1–10.

Vacquier, V. D., W. J. Swanson, and Y.-H. Lee. 1997. Positive Darwinian selection on two homologous fertilization proteins: what is the selective pressure driving their divergence? J. Mol. Evol. 44 (Suppl. 1): S15–S22.

Watanabe, J. M. 1982. Aspects of community organization in a temperate kelp forest habitat: factors influencing the bathymetric segregation of three species of herbivorous gastropods. Ph.D. diss., University of California, Berkeley.

Yamada, L., T. Saito, H. Taniguchi, H. Sawada, and Y. Harada. 2009. Comprehensive egg proteome of the ascidian *Ciona intestinalis* reveals gamete recognition molecules involved in self-sterility. J. Biol. Chem. 284:9402–9410.

Yang, Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24:1586–1591.

Yang, Z., S. Kumar, and M. Nei. 1995. A new method of inference of ancestral nucleotide and amino acid sequences. Genetics. 141:1641–1650.

Yang, Z., R. Nielsen, N. Goldman, and A. M. K. Pedersen. 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. Genetics 155:431–449.

Yang, Z., W. S. W. Wong, and R. Nielsen. 2005. Bayes empirical Bayes inference of amino acid sites under positive selection. Mol. Biol. Evol. 22:1107–1118.

Zigler, K. S., M. A. McCartney, D. R. Levitan, and H. A. Lessios. 2005. Sea urchin bindin divergence predicts gamete compatibility. Evolution. 59:2399–2404.

Zwickl, D. J. 2006. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. Ph.D. diss., The University of Texas at Austin.

Associate Editor: M. Hart